THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

Volume XXXVII February, 1946 Number 2

ON THE INTERPRETATION OF THE CORRELATION COEFFICIENT AS A MEASURE OF PREDICTIVE EFFICIENCY

HUBERT E. BROGDEN

Personnel Research Section, The Adjutant General's Office

It is the object of this paper to demonstrate that variation in efficiency of a selective or predictive instrument is a direct linear function of variation of its product-moment correlation with the predicted variable and to indicate possible criticisms of *E* ([3], p. 275) the coefficient of alienation, the coefficient of determination (r^2) and other functions of the correlation coefficient which purport to show efficiency of prediction. While the writer feels that most of the argument and interpretation presented is of general significance, caution should possibly advise limiting the discussion to those prediction problems most typical of psychological research, such as the prediction of success in a job, in an assignment, in training for industrial work, or in academic work. It is quite possible that the interpretation emphasized here has been presented elsewhere in the statistical literature. However, if this is the case, repetition would appear worth while in that current journals and statistical textbooks most often used by psychologists and educators usually present interpretations different from or contrary to that here developed.[1,2,3,4,5]

Since a number of the functions employed to interpret the correlation coefficient have their own legitimate meaning it may be well to indicate some typical comments which are in the author's opinion misleading. It is desired in indicating these comments to emphasize that interpretation of the correlation coefficient as a direct index of efficiency of a predictor definitely differs from interpretations usually advocated, and that the

practical consequences of these differences are appreciable and of practical importance. Thus Hull ([3], p. 275) has stated in speaking of the value of tests for industrial selection: " . . . forecasting efficiency rises as much (20 points) between correlations of .98 and 1.00 as between correlations of .00 and .60." In Guilford ([2], p. 364) we find: "Tests with a coefficient of validity less than .5 are practically useless except in distinguishing between extreme cases since the forecasting efficiency is only 13.4 per cent." In both instances E is employed as the means of interpreting the correlation coefficient. Peters and Van Voorhis ([5], p. 115), employing the coefficient of alienation ($\sqrt{1 - r^2}$) as their means of interpreting r, assert: " . . . where r equals .10, there remains 99.5 per cent of guess in a prediction based on it . . . " These statements are typical of those to be found in texts currently employed by psychologists.

In most prediction problems a test* or a battery of tests is employed as a means of selecting from a group of applicants those individuals who will perform most efficiently on the job. Hence some measure of efficiency on the job must be employed as a criterion. Presumably the criterion itself would be employed as a selector if expense, time, and effort were not important. It will be assumed here that the criterion is perfect and involves neither bias nor error of measurement. It is realized that it would be impossible to obtain such a criterion in practice. It might be noted, parenthetically, that it would, however, be possible to estimate the degree of error in the criterion and correct the correlation for attenuation of the criterion. Correction for error in the predictor would be undesirable, since evaluation of the predictor must take into account errors of measurement, and since such errors would be involved to exactly the same degree in the actual operating situation as in the experimental setup.

A measure of the efficiency of selection by the predictor should indicate directly the proportion of maximum saving that is actually obtained with use of the given test where the maximum saving is that obtained by selecting on the criterion. A zero value for the index of efficiency of prediction should be

* While the predictor will in general be referred to as singular, all discussion and comment will apply equally to a weighted or unweighted sum of several predictors or, in other words, to multiple and partial regression.

that which would be obtained with random selection; if one-half of the total possible saving were effected by use of the predictor, the obtained value for the index of predictive efficiency should be .5; while if one-fourth of the total possible saving were effected the index should be .25.

For any array of z_x in the correlation surface of z_x and z_y the z_y values can be obtained from the simple equation:

$$\bar{z}_y = r_{xy} z_x, \tag{1}$$

if the regression of z_x on z_y is linear. As in conventional notation, z_x and z_y indicate standard scores.* If men were selected as above a point of cut on the test scores, the relationship between their mean criterion and their test scores would be

$$\frac{1}{N} \Sigma \bar{z}_y = r_{xy} \frac{1}{N} \Sigma z_x, \tag{2}$$

where the N's and summations are within those applicants above the point of cut on the test. Equation (2) follows from (1) since the two summations obtained by adding a series of equalities and sums of equalities are equal to each other. Since r_{xy} is a constant in the right-hand member of (2) it may be factored out. It should be noted here that the distinction between $\Sigma \bar{z}_y$ and Σz_y disappears if N is large since errors of prediction are random and will 'average out.' Solving (2) for r we obtain

$$r_{xy} = \frac{\Sigma \bar{z}_y}{\Sigma z_x}. \tag{3}$$

If the distribution forms for both variables were the same, $\Sigma z_y'$ would be equal to and could be substituted for Σz_x. The necessity of this equality may be readily seen if it is remembered that equivalence of distribution form demands equivalence of

* The problem of absolute zero is implicitly ignored when variables are expressed in standard form, since the means are taken as zero points. This problem is not pertinent to that of selection except in setting critical scores and is definitely not related to the evaluation of the efficiency of a selective device. In evaluating efficiency of selection the problem is that of comparing random or average choice with that obtained by selecting with some measuring instrument. The problem of absolute zero would be of importance in selecting the test score array at which the average employee makes a contribution of zero; that is, the test score array where average cost of supervision and error balances the average value of output. The ramifications of the latter problem are many and will not be considered further in the present article.

frequency values for corresponding standard score values. Σz_x and Σz_y are both sums of products of frequencies and standard score values and must be equal to each other if the corresponding frequency and standard score values are equal. Substituting Σz_y for Σz_x, we obtain

$$r_{xy} = \frac{\Sigma \bar{z}_y}{\Sigma z_y'}, \qquad (4)$$

where $\Sigma z_y'$ indicates summation within a group selected on the 'criterion.' $\Sigma \bar{z}_y$ refers, still to a group selected on the test. Both groups are the same proportion of the total.

In (4) it can be seen that r is the ratio of the increase obtained by selecting above a given standard score on the test to the increase that would be obtained by selecting above the same standard score on the criterion itself. Thus, r, it would seem, is exactly the index desired for indicating the efficiency of a selective instrument. Equation (4) indicates that if an increase in mean criterion score (over the general mean) of Δ may be obtained by selecting above any given point on z_y, then an increase of $r\Delta$ may be obtained by selecting the same proportion of applicants by means of z_x. Note that an increase in validity from .00 to .10, from .35 to .45, or from .90 to 1.00 will produce exactly the same increase in the ratio of obtained to maximum goodness of the selected group if, of course, the percentage selected has remained constant.

Equation (1) may be readily converted into

$$r_{xy} = \frac{\bar{z}_y}{z_y'}. \qquad (5)$$

Thus in selecting a single case the correlation may still be interpreted as showing the ratio of the expected value of the gain effected by selecting with the test score to the gain that would be effected by selecting on the criterion itself.

While r_{xy} in (3) is quite meaningful in that it indicates the ratio of mean standard criterion scores to mean standard test scores of the selected group, the extension of the argument to (4) emphasizes the necessity of assuming agreement between the distribution of the criterion and that of the predictor before the product-moment coefficient may be properly regarded as an exact index of predictive efficiency no matter what z_x and z_y score is chosen as the cutting point for determining Σz_y and

$\Sigma z_y'$. It should possibly be pointed out here that if the distributions of the two correlated variables do not agree, the upper limit of the correlation between them is something less than plus one. Thus while in (3) it is probably generally true that an increase in correlation of .05 will mean an equally large increase in the mean standard criterion score of the selected group whether the correlation increase is from .20 to .25 or from .60 to .65; this relationship cannot hold for all cutting scores as r approaches 1.00 unless the frequency distributions of the two correlated variables are the same. (The correlation would still be a sort of average of the values obtained at the various cutting scores.)

When inequality between the frequency distributions of the predictor and the criterion is observed, the question arises as to the course of action indicated. Since this problem is at least indirectly related to that of selecting an index of predictive efficiency, we will digress for a moment and consider several points which seem to the author to be of some importance for prediction problems in general if not for the particular problem of this article. With such inequality it would be possible to alter either to obtain agreement with the second or to alter both to obtain agreement with some arbitrary standard, such as, for example, the normal curve. The writer believes that the criterion can, at least sometimes, be constructed so that units of the scale are equal at all levels and the resulting distribution form must consequently be accepted as it is empirically determined. Even though development of such criteria would rarely be possible, examination of the composition of such a criterion might help in improving actual criteria. Other arguments favoring various distribution forms seem hopelessly arbitrary.

Given the distribution form of the criterion, the test score distribution should then be such as to provide the 'best' prediction of the criterion. Probably the 'best' distribution of the predictor would closely resemble that of the criterion, especially if the correlation between them were high. However, in spite of the desirability of agreement from the viewpoint of simplicity in interpreting the correlation coefficient, there seems no basis for maintaining that the 'best' predictor distribution would be that which exactly agrees with the criterion distribution especially when the correlation is low. While the writer is of the opinion

that the distribution best for the predictor would tend probably to deviate from an anormal criterion distribution in the direction of normality, no attempt will be made here to demonstrate this point.

To obtain the desired equality of units of the criterion scale, some form of cost accounting procedures would be required. From the viewpoint of the industrialist or employee, the ideal criterion would indicate the cost to him of services or production units for each employee for the time interval during which the employee could be expected to work on the job for which he was selected. Since the essential comparison is between the selected employee and one chosen at random or at the average of the applicants in which selection is to occur, the criterion values would be conveniently expressed as mean deviates or as the difference in the cost of services by the given individual and the average such cost for a group representative of applicants in general. That is, if *A* produces 200 units in a given time interval as opposed to an average of 100, the cost to the employer of *A*'s produce would consist principally of *A*'s salary plus overhead on the equipment, office space, etc., used by *A*, while the cost to the employer of 200 units of produce by average employees would be twice that amount. Correction might be made here for spoilage of material, cost of errors or any other factors in which employees were found to differ, and which might be predicted by tests or measurements of abilities and other traits.

If the total cost in dollars were calculated for *A* and for the average employee the difference would constitute the saving to the employer, during the given time interval, effected by selecting *A* instead of choosing at random. If an estimate of the time interval *A* would be expected to remain on the job were available, a complete estimate of the monetary saving would be computed. This latter is the value proposed for use as a criterion. Units of such a scale would be comparable at all points to that the obtained frequency distribution of criterion values would have significance in its own right. Consideration of possible factors determining such a distribution of criterion scores suggests that they would tend often to deviate considerably from normality. Other things equal, 'spread' would be much greater at those class intervals on the distribution where individual differ-

ences represented considerable differences in cost of production or services, while at those class intervals where individual differences made for rather small differences in cost of production, 'bunching' would be apt to occur. Heavy skew or 'spread' would probably be found in criterion distributions for industrial jobs where infrequently occurring but costly errors or accidents must be considered in evaluating employees.

If the criterion were expressed directly in terms of dollars as just suggested, it might be considered desirable for some interpretative purposes to multiply both sides of (3) by the standard deviation of y. Since Σz_y would equal $\Sigma y/\sigma_y$, if y is in deviate form, (3) would become

$$r_{xy}\sigma_y = \frac{\Sigma y}{\Sigma z_x}. \tag{6}$$

The coefficient $r_{xy}\sigma_y$ does not indicate the efficiency of the predictor as a proportion of maximum saving and would not be appropriate for evaluating the test as such. It does, however, provide a direct estimate of the average saving in dollars per unit increase in the mean standard test score of the selected individuals. It would, of course, be possible to employ a coefficient such as r_{xy}/σ_x. The decision as to whether relationship between a variable should be expressed in raw or standard score form depends in part upon the degree to which raw scores are meaningful. The coefficient $r_{xy}\sigma_y$ would probably be more often useful in selection work than would r_{xy}/σ_x, since criterion scores are more likely to be directly meaningful than test scores.

In all the foregoing discussion of (3) and (4) and distribution forms, it has been assumed that the predictor is designed to select at all possible cutting scores. If, of course, it can be decided beforehand that the per cent to be eliminated is to remain at some fairly constant figure throughout the use of the test, it would be possible to concentrate items so as to select most efficiently [see Richardson[3] at that point]. The product-moment correlation would then be an inappropriate measure of predictive efficiency for such a test. In such instances a biserial correlation coefficient* could be computed after dichotomizing

* If the distribution of y is markedly anormal the problems associated with lack of agreement between the hypothetical normal x distributions and the obtained y distribution arise. While the only explicitly stated assumptions in arriving at the biserial formula is the assumption of linear

the test at the point of cut which would select the desired proportion of the population. From (3) the biserial may be obtained by substituting $N\ Z/P$ for Σz_x, where Z is the height of the ordinate of a normal curve at the point of cut and P the proportion of the population in the selected group. To obtain the conventional formula, $M_T - M_u/\sigma_T$ would then be substituted for $\Sigma z_y/N$; that is, $\Sigma z_y/N$ would be expressed in raw score terms where M_T and σ_T are the mean and standard deviation of the total group and M_u the mean y score in the selected or upper group.

DISCUSSION OF OTHER FUNCTIONS PURPORTING TO SHOW EFFICIENCY OF PREDICTION

The principal reason for preference of the correlation coefficient over other indexes which have been proposed as measures of predictive efficiency are positive in nature and depend upon the derivation developed in this article.

---

regression of y on x and that of normality of the x distribution, it can be seen that when r approaches unity these assumptions are inconsistent unless y is also normally distributed. The problems involved here are similar to those mentioned in discussing the necessity of assuming similarity between the distribution forms of x and y before the product-moment correlation could properly be regarded as an index of predictive efficiency. It is probably because of the inconsistency mentioned that biserial coefficients over plus one are often found in practice. Just as the correlation coefficient may be more clearly interpreted if the distribution forms are similar, so the biserial correlation would be subject to less misinterpretation if, instead of expressing the mean y value of the upper group in relation to the mean x value as it would be if it were normally distributed, it were expressed as it would be if x were distributed the same as y. The biserial would then be computed as in (4). The formula in raw score terms would be

$$r_{bis} = \frac{\left(\frac{M_y - M_u}{SD_y}\right)}{\left(\frac{M_y - M_{u'}}{SD_y}\right)} \qquad (a)$$

or, since the standard deviations would cancel,

$$r_{bis} = \frac{(M_y - M_u)}{(M_y - M_{u'})}. \qquad (b)$$

where the numbers in the selected groups are equal and M_u is computed on persons selected by means of the test, while $M_{u'}$ is computed on those selected by means of the criterion.

It appears to the author that much of the apparent problem in selection of the function of r which gives predictive efficiency in the most meaningful manner centers around the failure to realize that essentially the concern of selection is to increase the 'goodness' of the selected individuals, the failure to realize that the direct concern is with variation in z_y. It is only because of this concern that we are even indirectly interested in $(z_y - \bar{z}_y)$ or the error of prediction. Given this admission, it follows directly that goodness of the selected group is a function only of their test scores and the size of the correlation coefficient. No equation can be obtained in which the goodness of the selected group is a function of their mean test scores and the coefficient of alienation $(\sqrt{1 - r^2})$ or E $[100(1 - \sqrt{1 - r^2})]$. Both of these proposed indexes of predictive efficiency are concerned with error of prediction and are not directly pertinent to the primary concern of prediction.

Because of the nature of the relationship between r and $(1 - r^2)^{1/2}$, the distinction between them is highly important. The magnitude of r may reach nearly .44 before $(1 - r^2)^{1/2}$ falls below .90 and, conversely, r can reach .90 before $(1 - r^2)^{1/2}$ exceeds .44. Thus if the coefficient of alienation were taken as the means of interpreting the efficiency of a predictive instrument, tests having validities below .6 would be regarded as without practical value and improvement of tests with very high validity would be considered highly important. If, on the other hand, r is taken as the means of interpreting the efficiency of the predictive instrument, tests with validities of .3 or .4 or, in some instances, even tests with validities of .2, can still be considered useful, while the practical value of improving tests with validities above .9 would be considered negligible even if it were possible.

We had previously indicated that the interpretation of the correlation coefficient is the same no matter what the number of individuals in the population from which selection is made or the proportion in the selected group. To emphasize this point, transformation from equation (1) to (6) was made in order to indicate that even with a single case, the interpretation of the correlation coefficient is unchanged. It should be understood here that although the correlation coefficients is interpreted as a ratio of means as in (5), the correlation so interpreted is between individual cases and not between means. Apparently state-

ments frequently appearing in currently well-known textbooks to the effect that predictors with validities of .3 or .4 are of no use in making individual predictions but might be profitably employed to predict the performance of groups, are based on $\sqrt{1 - r^2}$ as the index of predictive efficiency for individuals, and $\sqrt{(1 - r^2)/N}$ as the index of predictive efficiency for groups. Both of these give the error of estimate in standard score terms. It should be clearly understood, however, that $\sqrt{(1 - r^2)/N}$ will give an estimate of how accurately we can estimate the efficiency of selection in a given sample and that it is only indirectly related to efficiency of selection itself. In other words, $\sqrt{(1 - r^2)/N}$ is analogous to an estimate of the standard error of a validity coefficient. As a matter of fact, if we consider (5) as a means of computing r instead of interpreting it, it is evident that the standard error of r as computed in this manner is some function of the standard error of Mz_y and Mz'_y where Mz_y at least would be determined by $\sqrt{(1 - r^2)/N}$. It would seem then that one would expect to gain as much per individual selected no matter whether one or one hundred individuals were selected so long as the basic population from which they were selected and the proportion of applicants selected is the same. It is true of course that testing per man would be somewhat more costly with very small groups and research prohibitively so. This, however, is not really pertinent to the point at issue.

Examples such as the following have often been cited as indicating a genuine difference between prediction for an individual and prediction for a group: Where an employer might be interested in the expected saving in money, or the head of a school in expected grades, the individual applicant might be more interested in whether he will 'make the grade' in a job and avoid being fired, or whether or not he will fail in school. This type of question might also be of primary concern to an individual offering vocational guidance. However, the difference in the statement of the problem here is not a function of the distinction between prediction for individuals and for groups, but of the nature of the predicted variable. The prediction problem in the example just cited differs from those previously discussed principally in that the criterion is regarded as dichotomous in the latter instance. Or, in other words, information as to whether the individual is above or below the point of cut is so

much more important than the degree to which he is above or below, that the latter information is considered negligible in relation to the former. Except in such special cases, it would seem in general undesirable to interpret a correlation in terms of a dichotomous criterion instead of a continuous variable. In addition to the reduction in the amount of information made available, the interpretation becomes more complex since the proportion of cases in the upper group bears a curvelinear relation to the predictor. Even given the proportion in the upper group—or the probability that the individual will fall in the upper group—the interpretation is no simpler than that obtained by determining the comparable figure for the continuous variable (z_y), except possibly in dealing with individuals sufficiently unsophisticated in statistics that the concept of standard scores would not be understood.

The apparent advantage of r^2, the coefficient of determination, is its legitimate interpretation as the proportion of criterion variance accounted for by the predictor. This advantage is most evident in dealing with uncorrelated predictors since, in such instances, the sum of their coefficients of determination is equal to the square of their multiple correlation. However, the fact that r^2 is additive under certain conditions does not mean that it is a measure of efficiency of prediction. While it is advantageous in many mathematical manipulations and corresponding arithmetic calculations to employ an index which is additive, it is not relevant to the choice of an index of efficiency of prediction. The reader can readily convince himself by attempting algebraic manipulations of (1) that no equation can be obtained in which r^2 can be directly employed for estimating the efficiency of the performance of the selected group. The only meaningful expression is obtained by squaring both sides of (1), summating over the entire population and dividing by N. Such manipulations will demonstrate only that the variance of z_y is equal to r^2.

The interpretation of the correlation coefficient in terms of common elements does not appear directly relevant to the present discussion, since it is not usually proposed as a means of interpreting correlations as a measure of the efficiency of selective instruments. Such interpretation indicates, as the reader is probably aware, that under certain conditions r_{xy} can be taken

as the proportion of hypothetical constituent elements common to the variables x and y. However, Guilford ([2], p. 365) further indicates that when all x elements are also y elements, the percentage of such y or criterion elements determined (or the percentage of such elements in common) is then equal to r^2. This latter point is mentioned as a basis for interpreting r^2 as the coefficient of determination. This interpretation is analogous to the interpretation of r^2 as the percentage of variance accounted for by the predictor. There seem to be no additional factors involved here that would justify modification of the comments on r^2 made in the preceding paragraph.

SUMMARY

It has been shown that if the regression of y on x is linear and their frequency distributions the same, the correlation coefficient gives directly the ratio of the mean standard criterion score of a group selected by means of the predictor to that which would be obtained by selecting a group of similar size by means of the criterion itself. This is interpreted as showing that the correlation coefficient is a direct index of predictive efficiency. Related problems and other proposed indexes of predictive efficiency have been discussed.

REFERENCES

1) Garrett, Henry E., *Statistics in Psychology and Education;* 1944, pp. 493.

2) Guilford, J. P., *Psychometric Methods;* 1936, pp. 535.

3) Hull, Clark L., *Aptitude Testing;* 1928, pp. 535.

4) Hull, Clark L. "The correlation coefficient and its prognostic significance." *J. Educ. Res.*, 15, 327–338.

5) Peters, Charles C. and Van Voorhis, Walter R., *Statistical Procedures and Their Mathematical Bases;* 1940, pp. 516.

6) Richardson, Marion W. "The relation of difficulty to the differential validity of a test." *Psychometrika,* 1936, 1, 33–49.

MOTIVATION IN OCCUPATIONAL CHOICE AMONG JUNIOR-SENIOR HIGH SCHOOL STUDENTS

URBAN H. FLEEGE

The Catholic University of America

AND

HELEN J. MALONE

Assistant to the Director, American Junior Red Cross

If we are to help our young people find their place in life we must ascertain not only what they as individuals are capable of doing and what they are interested in doing, but why they wish to do the things they want to do. We must know the motives prompting their vocational choice if the guidance we give them in helping them formulate their life plan is to be intelligent and of lasting value.

Interest and ability are only two of the three important factors that make for success and happiness in a chosen vocation. Sound motives are basic to persistent effort in any vocation and lie at the very root of satisfaction and sense of achievement which one's work must yield if happiness is to be the result of occupational endeavor.

For this reason a vocational guidance program should be motive-centered. Motives play a major rôle in everyone's life; consequently, our school program must be conscious of the inculcation of motives. In one form or another educators state that one of the major objectives of our schools is the development of character. But the background of a program aiming to develop character is essentially the inculcation of motives, for the dominant characteristic of a man of character is strength of will, but will power is basically motive power.

An important question then, is, to what extent are our schools developing motives? What kind of motives are taking root in the minds of our adolescents? More specifically, what values influence the young adolescent's choice when he selects a life vocation?

Such questions prompted the investigation of the motives for vocational choice of five hundred thirty-three boys and girls from two junior-senior public high schools in New York State,

one a suburban school with a registration of six hundred fifty and the other a rural central school of two hundred pupils.

Through the English teachers, the pupils were asked to write a composition on the subject, "What I Want to be When I Leave School After the War and Why I Chose This Occupation." Compositions from two hundred twenty-two boys and three hundred eleven girls ranging in age from eleven to nineteen, representing grades VII through XII were analyzed to discover occupations and motives according to grade, sex, age, and IQ.

ANALYSIS OF OCCUPATIONS CHOSEN

An analysis of occupations chosen reveals that fifty-three per cent of the boys and forty-seven per cent of the girls chose careers in the professional and managerial fields, indicating a close correlation with the findings of similar studies.[1] Such choices ill accord with the actual opportunities available in these careers; consequently, the need of intelligent vocational guidance. For example, Bingham found that out of every thousand workers only five were actively engaged as physicians, whereas out of one thousand students one hundred and five wish to be doctors.[2]

In the professional field nursing was the most popular choice with the girls, accounting for thirty-nine per cent of all professional occupations chosen; whereas with the boys the most popular occupation was engineering; in nearly every case they mention a specific branch of engineering. Roeber and Garfield found the same to be true in their study.[3] There is no doubt but that the recent national drive for more nurses and engineers was partly responsible for the popularity of these occupations among the subjects of our study.

Second and third choices in the professional group among the boys were aviation and "running a business." Among the

---

[1] Urban H. Fleege, *Self-Revelation of the Adolescent Boy.* Milwaukee: The Bruce Publishing Co., 1945, 98.

[2] Walter V. Bingham, *Aptitudes and Aptitude Testing.* New York: Harper and Bros., Revised Edition, 1937, 108.

[3] Edward Roeber and Sol Garfield, "Study of Occupational Interests of High School Students in Terms of Grade Placement," *Journal of Educational Psychology,* 34: 355, October 1943.

girls other occupational choices included teaching (seventeen per cent), costume designing (two and one-half per cent), and a few scattered choices including such specialized branches as stylist, public welfare investigator, veterans' counselor.

The second highest group of occupations chosen by the girls was in the clerical field, twenty-nine per cent choosing vocations in this area. The most popular job was that of secretary (fifteen per cent) while stenographer and bookkeeper ranked second and third. Contrasted with the girls are the few boys (three per cent) selecting a clerical position. The group of occupations meriting second place among the boys was the skilled trades; we found one out of every five boys electing to become a mechanic; by contrast we found the girls showing no interest whatever in the skilled trades. This same divergence of choice was evident in the agricultural field, where we found thirteen per cent of the boys choosing farming or forestry as a career, to rank this occupation third in popularity, as compared with a complete lack of interest in this field among the girls. This appears strange in view of the fact that thirty per cent of the girls included in our investigation were from a distinctly rural area.

Among the service occupations which ranked third with the girls, beauty culture and airline hostess were the outstanding choices. Among the boys, service occupations ranked fourth; in this field careers in the army and navy were popular, but principally with the younger boys.

Very few girls (three per cent) and still fewer boys (two per cent) chose semi-skilled or unskilled occupations. Only one per cent of our entire group chose factory work despite the fact that a fair percentage of them live in an industrial area. These findings present quite clearly the problem facing high school counselors. A very large percentage of workers earn their living in the semi-skilled and unskilled trades, yet only two or three high school students out of a hundred favor such occupations.

GRADE LEVEL AND IQ AS FACTORS IN OCCUPATIONAL CHOICE

Considering the choice of occupation from the standpoint of grade level, we note a decrease in choice of occupation on the professional level as we advance from grade VII to XII. This trend is more in evidence among the girls, among whom fifty-six

per cent of the seventh-graders chose a professional career as compared with forty per cent of the twelfth-graders, than it is among the boys, among whom forty-eight per cent of those in the seventh grade chose a profession as compared with forty per cent in grade XII. In the clerical occupations the girls give evidence of a decided increase of interest as they grow older; this is particularly noticeable in the tenth and again in the twelfth grade. Perhaps this is partly because in New York State it is rather difficult for students to take commercial subjects before the tenth year. The increase in clerical choices in the final year may be due to economic reasons and the sudden relinquishing of hope to go to college. No definite age trends were evident in the other vocational preferences.

Examining occupational choice in the light of native intelligence as reflected by the IQ, we find little correlation between ability and the intellectual challenge of an occupation. On the contrary, we find evidence of the idealism of adolescents and their desire to live among the stars. Among those choosing professional occupations we find IQs' ranging from 65 to 156 among the girls and from 65 to 145 among the boys. Among both boys and girls who chose a profession we find a slight indication of the effect of maturity on occupational choice, as only ten to fifteen per cent of the eleventh- and twelfth-graders whose IQ range was below 95 chose a professional occupation, whereas in the lower grades, between twenty and forty-five per cent of those who made a similar choice had an IQ below 95. There is almost an exact duplication of these findings among the girls choosing clerical occupations.

Analyzing the occupational choices in the upper range of IQs 115 and above, we find less than one per cent of the girls and less than six per cent of the boys choosing an occupation outside the professional or clerical group; the majority of boys who did so choose selected a skilled occupation. In general, there is a slight tendency for the majority of students selecting agricultural, service, semi-skilled, and unskilled occupations to come from the lower IQ range, but this trend is by no means uniform. An analysis of our data suggests that both boys and girls, at all grade levels, tend to choose occupations without sufficient consideration of their ability to do the work of their choice. In this point our data corroborate the findings of Lehman and

Witty who noted little change in efficiency of vocational choice with advancement in chronological age.[4]

MOTIVES FOR CHOOSING A PARTICULAR OCCUPATION

The motives which the boys and girls of our study advance for a particular occupational choice are dominantly and almost exclusively on the purely natural plane. Of the six hundred eighty-nine motives given, only two could be classified as religious or supernatural. As is evident from Table I, less than a sixth of all the motives offered can be classified as motives of the 'highest natural level.' More than seventy-five per cent of the reasons advanced for choosing a particular vocation are selfish. Slightly less than seven per cent were unable to think of a reason for their occupational choice.

TABLE I.—MOTIVES FOR OCCUPATIONAL CHOICE DISTRIBUTED ACCORDING TO VARIOUS LEVELS OF MOTIVATION

Levels of Motivation	Number of boys and girls advancing motive
1. Religious or supernatural motive	2
2. Superior natural motive	111
a) Parental love	0
b) Filial love	6
c) Fraternal love	83
d) Romantic love	22
3. Other unselfish natural motives	20
4. Selfish natural motives	520
5. No motive	36

In Table II we note how interest in work was the most frequently mentioned motive. This reason was offered by more than a third of the boys and girls and accounts for twenty-eight per cent of all the motives listed. Help to mankind—the outstanding unselfish motive advanced—gives evidence of a striking sex differentiation. Of all the students offering this humanitarian motive, seventy-seven per cent were girls. A factor partly responsible for this high percentage is the large number of girls who chose nursing as a career. The only other motives

[4] Lehman and Witty, "Vocational Guidance, Some Basic Considerations," *Journal of Educational Sociology*, November, 1934, 174.

which the girls expressed much more frequently than the boys were: desire to meet people and advance socially, mentioned by four times as many girls as boys; desire for enjoyment and happiness, listed by twice as many girls as boys; and desire for home and children, a motive expressed only by the girls.

TABLE II.—DISTRIBUTION OF MOTIVES ADVANCED FOR CHOOSING A PARTICULAR OCCUPATION

Motive	Percentage mentioning motive	
	Boys	Girls
1. Interested in work	36	37
2. Want to help mankind	9	22
3. Personal advancement	19	11
4. Aptitude, feel fitted for job	11	10
5. Adventure	11	5
6. Occupation of family members	11	4
7. Favorable work conditions	7	6
8. Desire for home and children	0	9
9. Enjoyment, happiness	3	6
10. Security and Independence	5	3
11. Opportunity for work	6	1
12. Social advancement	1	5
13. To get work experience	4	1
14. Persuasion	1	2
15. Gratitude to family	1	1
16. Love of nature	1	0
17. Desire for fame, love of God, easy work	.5	.5
18. No motive	7	7

Motives which the boys mentioned twice as often as the girls were: opportunity for personal advancement, desire for adventure, desire to follow in father's footsteps, and a desire for security and independence. The motives, opportunity for work and work experience, were mentioned almost exclusively by boys.

Interest in work, a feeling of possessing an aptitude for a particular job, and favorable work conditions, were the only motives among those more frequently mentioned in which there appeared a close similarity of reasons for occupational choice among boys and girls.

INFLUENCE OF GRADE LEVEL AND IQ ON MOTIVATION

Analyzing our data according to grade level, there seems to be little difference in the level of motives of students from the seventh to the twelfth grade. No age trend seems evident; for example, as many of the younger (roughly ten per cent) as the older students mention the humanitarian motive, 'help mankind,' as the reason for their choice of occupation. The only exception to this lack of change or development of motives through the junior-senior high school years occurs in the motive—aptitude for work—in which we find a steady increase from four per cent among the seventh-graders to thirteen per cent among the seniors. Perhaps this growth in appreciation of aptitude for an occupation can be attributed to the work of the part-time guidance counselor who services the schools which these students attend.

In studying our classification of motives in the light of IQ variation we note no significant trend for those of higher intelligence to offer superior motives for occupational choice, nor for those of lower IQ to offer less noble motives.

RELATION BETWEEN MOTIVE AND TYPE OF OCCUPATION CHOSEN

When we analyze our data to discover the relationship between the type of occupation chosen and the motive given for the choice, we find that fifty-one per cent of all the motives given in our study were advanced by the fifty per cent who chose professional careers. Twenty-two per cent of all the motives given were offered by those choosing clerical occupations; six per cent by those choosing service occupations, six per cent by those choosing agriculture, nine per cent by those choosing semi-skilled trades, and four per cent by those choosing unskilled work.

The entire range of motives is represented by those choosing professional careers, but the majority (seventy-four per cent) of their motives cluster about the four most popular motives: interest in work (thirty-one per cent), help to mankind (twenty per cent), personal advancement (twelve per cent), and aptitude for the work (eleven per cent).

Among those choosing clerical occupations, thirty-nine per cent give interest in the work as their motive; thirteen per cent mention personal advancement; six per cent, aptitude. Thus

the four top-ranking motives in our study motivate less than six out of ten of the prospective clerical workers. Desirable working conditions and social opportunity were among the additional motives more frequently mentioned.

Students who selected service occupations, (e.g. airline hostess), mentioned adventure and travel as their dominant motives, while those choosing agriculture mentioned their love of outdoor life, working conditions, and the fact that farming is a family occupation. Of those choosing the skilled trades, fifty-one per cent were motivated by interest in the work, eighteen per cent were looking for personal advancement, and fourteen per cent chose this type of occupation because their father or brothers were in this field. Interest in the work, security, and personal advancement were the only motives given by the few students choosing the semi-skilled and unskilled trades. In general we may conclude that motivation for professional work is on a higher level than for the other job classifications.

It is striking that both the boys and the girls regard all occupations other than professional careers, as just mere ways of earning a livelihood. They do not seem to grasp the true relationship between the various occupations and society at large. Their attitudes give little evidence that they understand the contribution which the various occupations make toward personal and social welfare. Our findings suggest that one of the big tasks for our teachers and guidance counselors is to develop proper attitudes toward the non-professional vocations. Youth must realize not only the necessity of non-professional workers, but the important contribution that can be made to the world through the non-professional occupations.

To give a more intimate picture of the reaction of students to the question of motives in vocational choice, we present several cases.

Case I. A. S., male, grade VII, age twelve, IQ 102, nationality background: English, chosen career: scientist.

Motivation in the case of A. S. is very strong. Several members of his family, as well as himself, are victims of the bone disease, *osteogenesis imperfecto*, and his one desire is to work on a cure for this disease in order to help not only his own family, but all others who may be victims. In his own words, "My first problem will be to try to make cripples walk again." He

feels that his own experience will furnish him the motive to overcome the hardships of the task he has set for himself.

Case II. A. P., male, grade IX, age fourteen, IQ 115, nationality background: Italian, chosen career: radio work.

A. P. says, "I have chosen radio because of its interesting work and great future." His motive is on the natural level; he goes on in his composition to explain that in the postwar world radio and television will be increasingly important and that he would like to invent a very small radio which, as he says, "would set a man up for life."

Case III. S. A., female, grade XII, age seventeen, IQ 101, nationality background: Austrian, chosen career: journalism.

S. A. chooses journalism because "journalism seems to bring me in the field of working directly with the general public, and extending my efforts and ability in seeking information that in some form may be a benefit to all who need it."

Case IV. P. B., male, grade VIII, age thirteen, IQ 142, nationality background: American, chosen career: mechanical engineering.

"There will be so much to be done to return our country into a busy peace-time industrial center . . . There seems to be an endless task ahead for one who would be interested in building for the future and making our world one where anyone who wants to can improve his way of living and be given a chance to work out his own ideas for a better place to live."

Case V. L. A., female, grade X, age fifteen, IQ 87, nationality background: English, chosen career: marriage.

"I don't think I would be very successful in the business world and at present have no idea what I shall do after I graduate. I hope to get married some day and live in a little house outside of town where we can have a garden, chickens, horses and dogs. I have learned to do some cooking and housework, and can care for a baby; so if I do work after I leave school I'd rather do something along this line."

SUMMARY AND CONCLUSIONS

Of the five hundred thirty-three students in our study, fifty per cent chose professional or managerial occupations. Our findings give evidence of a definite need for intelligent guidance in vocational choice and for the development of a better under-

standing of the importance of the other-than-professional occupations.

The clerical field appeals primarily to the girls, while the skilled trades appeal to the boys.

Both boys and girls, especially at senior-high level, evinced a fair knowledge of occupations, choosing a total of one hundred seven different jobs, the boys seventy-eight, the girls fifty.

A large percentage of adolescents chose occupations beyond their mental capacities, thus giving decided evidence of their youthful idealism.

The motives advanced for deciding on a particular vocation were almost exclusively on the natural level; approximately thirty per cent did not go beyond interest in the work; seventy-five per cent of the motives mentioned were selfish; only two students mentioned a religious motive. These findings seem to point to the lack of higher motives in the adolescent's thinking and indicate the dire need of religion in the adolescent's life.

There was little change in motive from grade to grade. It would seem that the schools have either neglected the inculcation of higher values, or have failed in their efforts to develop higher motives as an influence in important life decisions.

There seemed to be no relation between the range of IQ and the motive advanced for a particular vocational choice.

Sex seems to be a factor in regard to certain motives. Certain reasons are mentioned more frequently by boys, others by girls. Unselfish motives are more influential with girls than with boys.

The four most popular motives that influenced the majority of students who chose a professional or clerical career as well as those who chose another occupation were: interest in work, opportunity to help mankind, opportunity for personal advancement, and aptitude for the work.

Our findings seem to point quite definitely to the need for the inculcation of an appreciation of higher values, particularly religious values, on the junior-senior high school level. Adolescents must be lead to grasp and personalize motives and ideals which will lift their vision beyond the selfish, the material, and the natural.[5] They must be made aware that a vocation is more than an occupation—that it is the means of fulfilling their life purpose.

[5] *Self-Revelation of the Adolescent Boy*, 328.

THE EFFECT OF PRE-FIRST-GRADE TRAINING UPON READING READINESS AND READING ACHIEVEMENT AMONG SPANISH-AMERICAN CHILDREN

SELMA E. HERR

College of Education, University of New Mexico

Many studies have been made of the Spanish-American children in Arizona, California, New Mexico, and Texas. The investigators have appeared to conclude that the Spanish-American children are a special problem in our schools and that they need to be handled by a method other than that used with English-American children. These writers have generally recognized a language handicap and some have discussed the influence of environmental conditions and past experience.

PURPOSE OF THIS STUDY

The purpose of this study was to find the effect of a year of pre-first-grade training upon the reading readiness and achievement of Spanish-speaking children in the first grade. If Spanish-speaking children are taken into the schools a year in advance of the time when they normally enter, what advantage will they have over others who enter at the usual time?

The question is one of practical importance because of the large retardation of Spanish-speaking children in the first grade under the usual school policies. The delay in learning to read, results in an excessive overage condition with its attendant evils. If it is discovered that a considerable part of the handicap with which these children start can be overcome before the normal time for introducing a reading program, the educational implications will be significant.

PROCEDURE

Nine towns in New Mexico were selected for the experiment. In August, 1940, through the use of the Pintner-Cunningham Primary Intelligence Test, Form B, and by a subjective analysis of vocabulary ability and home environment, two groups of five-year-olds of these communities were equated for the control and the experimental groups.

During the year of 1940, the control group did not attend school and were subjected to no special training. The experimental group attended school and were subjected to a pre-first-grade curriculum directed toward:

1) Social and emotional adjustment through democratic procedures in conversation, discussion, acquaintance with the school organization, room responsibility, and planned games and parties.

2) Vocabulary development through games, discussions, conversations, listening, and describing pictures, books, and objects.

3) Physical development through games, craft media, rhythmic and creative play, and nutritional and health habit formation.

4) Development of auditory perception through listening to sentences, poems, stories, and elements in phonograph records.

5) Development of visual perception through looking at pictures and observing similarities and differences in objects.

6) Development of habits of memory through listening to nursery rhymes, learning to carry messages correctly, and reproducing of stories.

7) Development of coöperativeness and social attitudes through the correlation of the above mentioned procedures.

In order to help the child in his orientation to the school environment, the school atmosphere was made as pleasant and homelike as possible.

In the fall of 1941, both the control and the experimental groups began the usual first-grade program. They were given the Pintner-Cunningham Primary Intelligence Test, Form A, and the Metropolitan Reading Readiness Test in November. The results for the above named tests, together with the results of the Pintner-Cunningham Primary Intelligence Test, Form B, are as follows:

Test		Experimental Group	Control Group	Date
Pintner-Cunningham, Form B	Mean	66.01 IQ	68.9 IQ	Aug. '40
	Sigma	7.4	7.14	
Pintner-Cunningham, Form A	Mean	95.92 IQ	78.46 IQ	Nov. '41
	Sigma	7.3	11.9	
Reading Readiness	Mean	52.0	7.39	Nov. '41
	Sigma	6.18	5.61	

Since the tests were administered in English, and since the children had only a limited English vocabulary, the results serve merely as a basis for comparison and diagnosis, and not as a true measures of intelligence.

In the beginning of the second year of the experiment, a reading-readiness program was presented in order to promote growth in various abilities and interests of reading maturation.

The reading experiences were based upon the actual experience of the child in order to develop his interest and to stimulate his imagination. The reading experiences were meaningful processes realized through lifelike situations familiar to the child. The three major tasks of the reading program were the development of ideas through an enrichment of experience, the development of a meaningful vocabulary to express the ideas gained through these experiences, and the development of ability to recognize printed symbols which express, in printed form, these ideas and meanings.

The first reading was introduced through the use of the experience charts which were based upon the child's actual experience. When the pupil showed a definite desire to read, the preprimer was placed in his hands. There were modifications or differences in the program for the two groups based upon their abilities.

In May, 1942, the children of both groups were given the Metropolitan Achievement Test: Primary I Battery, Form B.

TEST RESULTS AND THEIR SIGNIFICANCE

As previously stated, this study is concerned with first-grade Spanish-American children, half of whom had received a year of pre-first-grade training and half of whom had not had such training. The Table I presents an outline of the testing program used in the experiment.

Through the statistical treatment of these test results it was expected that certain questions pertaining to the study could be satisfactorily answered. As stated, the specific purpose of the experiment was to determine the effect of a pre-first-grade training program upon scores made in reading readiness and reading achievement tests by Spanish-American first-grade children. The groups were matched in August, 1940, on the

TABLE I.—TESTING PROGRAM

Test	Date	Number of Children
Pintner Cunningham Primary Intelligence Form B	August, 1940	469
Pintner Cunningham Primary Intelligence, Form A	November 14–18, 1941	301
Metropolitan Reading Reading Readiness	November 21–25, 1941	301
Metropolitan Achievement Primary I Battery: Form B	May 13–17, 1942	200

TABLE II.—CHRONOLOGICAL AGES OF CHILDREN (NOVEMBER, 1941.)

CA in Months	Number of Children	
	Experimental	Control
105–108	0	1
102–104	2	3
99–101	2	0
96–98	0	1
93–95	1	2
90–92	0	4
87–89	2	6
84–86	9	18
81–83	9	12
78–80	25	16
75–77	14	20
72–74	31	10
69–71	2	7
66–68	3	0
Total	100	100

basis of performance on the Pintner Cunningham Primary Intelligence Test, Form B, and on home environment and language ability. The chronological ages of the children are given in Table II.

It should be noted that the average age of the pupils who had pre-first-grade training was five years and three months in September, 1940; that is to say, for the most part, they were

TABLE III.—IQ RATINGS AS OBTAINED FROM THE PINTNER CUNNINGHAM INTELLIGENCE TEST, FORM B.
(AUGUST, 1940.)

IQ	Number of Children	
	Experimental	Control
78–80	1	5
75–77	24	35
72–74	13	8
69–71	9	4
66–68	1	10
63–65	7	18
60–62	16	11
57–59	24	10
54–56	5	0
Total	100	100
Mean	66.01	68.9
Sigma	7.4	7.14

children who would not normally be enrolled in school. The normal chronological age for first-grade children, assuming they enter school in September at the age of seventy-five months, would be seventy-eight months in November. The experimental group used in this project was almost a month younger than the average age, while the control group was almost three months older chronologically than the average normal first-grade child. The means for the two groups were:

	Experimental Group	Control Group
Mean	77.40 months	80.88 months
Sigma	6.72	7.62

It will be noted that the children of the experimental group were three and a half months younger than those of the control

TABLE IV.—IQ RATING AS OBTAINED FROM THE PINTNER CUNNINGHAM INTELLIGENCE TEST, FORM A. (NOVEMBER, 1941.)

IQ	Number of Children	
	Experimental	Control
135–139	1	0
130–134	3	0
125–129	5	0
120–124	3	0
115–119	7	0
110–114	8	0
105–109	9	0
100–104	11	2
95–99	17	8
90–94	16	2
85–89	20	6
80–84	0	11
75–79	0	39
70–74	0	26
65–69	0	6
Total	100	100
Mean	95.92	78.46
Sigma	7.3	11.9

group. Later results from the reading readiness and reading achievement tests indicate that the chronological age within the age range of the children in this study showed no sign of influence on reading readiness and achievement.

Table III compares the IQ's of the experimental and control groups as obtained from the Pintner Cunningham Primary Intelligence Test, Form B, as given in English, August, 1940.

Tests administered to the Spanish-speaking child are not always true measures of intelligence because the tests usually have been standardized upon the children who have had a different cultural and language background. That the Spanish-American children have been considered inferior in intelligence because of these tests, is unfortunate.

TABLE V.—MENTAL AGES OF CHILDREN AS OBTAINED FROM PINTNER CUNNINGHAM TEST, FORM B TEST SCORES. (AUGUST, 1940.)

Mental Age in Months	Number of Children	
	Experimental	Control
69–71	1	0
66–68	1	1
63–65	2	2
60–62	5	1
57–59	8	2
54–56	27	25
51–53	20	15
48–50	18	12
45–47	8	10
42–44	6	9
39–41	3	13
36–38	1	0
Total	100	100
Mean	52.62	53.55

The Pintner Cunningham Primary Intelligence Test, since it was given in English to children of limited English-speaking vocabulary, was not a valid measure of these children's intelligence and was used merely as a tool for equating the control and experimental groups. The Table IV shows the increase in performance on the test fifteen months later.

The IQ's were raised 29.91 points for the experimental group after a year of pre-first-grade training. The IQ's for the control group were raised from 68.90 to 78.46; these pupils had been in school two months at the time the second test was given.

TABLE VI.—MENTAL AGES OF CHILDREN AS OBTAINED FROM THE PINTNER CUNNINGHAM PRIMARY TEST, FORM A, TEST SCORES IN NOVEMBER, 1941.

Mental Age in Months	Number of Children	
	Experimental	Control
114–116	1	0
111–113	0	0
108–110	0	0
105–107	1	0
102–104	3	0
99–101	4	0
96–98	3	0
93–95	2	0
90–92	4	0
87–89	4	0
84–86	4	0
81–83	7	5
78–80	5	1
75–77	5	8
72–74	7	2
69–71	12	10
66–68	23	10
63–65	14	13
60–62	11	20
57–59	0	8
54–56	0	20
51–53	0	3
Total	100	100
Mean	76.83	64.05

The mental ages as obtained from the Pintner Cunningham Primary Intelligence Test, Form B, as administered in English, in August 1940, are shown in Table V.

In August, 1940, the mean mental age for the experimental group was 52.62, and for the control group, 53.55. It should be repeated that these are not true mental ages, but mental ages of children measured by a test in a language with which most were relatively unfamiliar. Table VI gives the mental ages of these same children in November, 1941, as obtained from

TABLE VII.—PERCENTILE RANKS ON READING READINESS TEST NORMS (NOVEMBER, 1941)

PR	Number of Children	
	Experimental	Control
90–95	1	0
85–89	2	0
80–84	3	0
75–79	4	0
70–74	6	0
65–59	0	0
60–64	6	0
55–59	16	0
50–54	13	0
45–49	16	0
40–44	12	0
35–39	8	0
30–34	6	0
25–29	4	7
20–24	6	11
15–19	7	11
10–14	0	37
5–9	0	20
0–4	0	14
Total	100	100

Form A of the Pintner Cunningham Tests, after the experimental group had been in school a year and three months, and the control group had been in school for three months.

Table VI shows that the average mental age of the experimental group in November, 1941 was 1.83 months above the normal mental age for first-grade children. These children had increased their average mental age 24.21 months since August, 1940. The average mental age of the control group was 64.05 or ten months below the average normal mental age of the first grade.

Table VII gives the percentile ranks on the Metropolitan Reading Readiness Test norms.

Using these percentile ranks as a basis of prediction in attaining a given grade status at the end of the first grade, Table VIII shows the placement could be predicted.

TABLE VIII.—PREDICTION OF GRADE STATUS AT END OF FIRST GRADE (NOVEMBER, 1941)

Grade Prediction	Number of Children	
	Experimental	Control
1–9 or over	16	0
1–8	6	0
1–7	18	0
1–6	13	0
1–5	7	0
1–4	12	0
1–3	6	0
1–2	7	0
1–1	7	10
1–0	8	87
Total	100	100

The figures in Table VIII indicate, in the light of prediction of grade status at the end of the first grade, that sixteen per cent of the experimental group will attain the grade level of grade

one and nine months, and over fifty per cent will attain a grade level of grade one and six months. All the children in the control group were below the grade attainment of grade one and three months.

The administration of the Metropolitan Achievement Test: Primary I Battery, Form B, in May 1942, indicated the grade placement given in Table IX.

TABLE IX.—READING GRADE PLACEMENT AS OBTAINED FROM THE METROPOLITAN READING ACHIEVEMENT TESTS (MAY, 1942)

Grade Placement	Number of Children	
	Experimental	Control
2–3	3	0
2–2	5	0
2–1	12	0
1–10	18	0
1–9	14	0
1–8	27	0
1–7	17	0
1–6	4	0
1–5	0	4
1–4	0	8
1–3	0	22
1–2	0	47
1–1	0	19
Total	100	100

Table IX shows that twenty per cent of the experimental group had achieved a grade placement of 2–1 or above by May, 1942. Four children had a grade placement lower than 1–6. In the control group, the highest grade achievement was 1–5.

In referring to the prediction of attainment as obtained from the percentile ranks on the Metropolitan Reading Readiness Test, given in November, 1941, it will be noted that it was pre-

dicted that sixteen per cent would attain the level of grade 1–9, but fifty-two per cent of the children in the experimental group actually attained this level. The Achievement scores of the control group indicate that four per cent of the children had attained the grade placement of 1–5, and eighty-eight per cent were at or below the grade placement level of 1–3. In referring

TABLE X.—READING ACHIEVEMENT AGE AS OBTAINED FROM THE METROPOLITAN READING ACHIEVEMENT TEST (MAY, 1942)

Achievement Age in Months	Number of Children	
	Experimental	Control
87	1	0
86	4	0
85	4	0
84	11	0
83	0	0
82	1	0
81	14	0
80	16	0
79	30	0
78	16	0
77	3	0
76	0	4
75	0	8
74	0	23
73	0	43
72	0	22
Total	100	100
Mean	80.9	74.1

to Table VIII, it will be noted that eighty-seven per cent of the control group were expected to have a grade placement of 1–0, ten per cent a level of 1–1, and three per cent a level of 1–2.

The reading achievement ages attained by the scores of the Metropolitan Achievement Test are given in Table X.

In November, 1941, the experimental group, through the interpretation of the scores of the Pintner-Cunningham Primary Test, Form A, had an average mental age of 76.83 months. In May, 1942—six months later—they had attained an average reading achievement age of 80.9 months on the Metropolitan Achievement Tests. For the same period of time, the control group with a mental age of 64.5 in November had attained an average reading achievement age of 74.1 months.

TABLE XI.—ACHIEVEMENT AGE IN ARITHMETIC AS OBTAINED FROM THE METROPOLITAN ACHIEVEMENT TEST (MAY, 1942)

Achievement Age in Months	Number of Children	
	Experimental	Control
87	5	0
86	7	0
85	4	0
84	7	0
83	0	0
82	0	0
81	14	1
80	22	3
79	14	1
78	7	2
77	2	4
76	1	3
75	8	5
74	5	13
73	4	25
72	0	43
Total	100	100
Mean	80.5	74.0

In the experimental group, fifty-one per cent of the children attained a reading achievement level equal to or above that of normal first-grade children, as set forth by the norms of the

Metropolitan Achievement Test. None of the control group had reached the average norm for first-grade reading achievement.

The arithmetic scores parallel those in reading as shown in Table XI.

The mean achievement age in arithmetic for the experimental age was 80.5 months as compared to 80.9 months in reading achievement. The mean achievement age in arithmetic for the control group was 74.0 months as compared to 74.1 months in reading achievement.

In both instances, the experimental group grew mentally more than two thirds of a school year more than did the control group in the six months period from November to May.

CONCLUSIONS

1) The average IQ, however—not the true IQ—of the experimental group as obtained from the scores of the Pintner-Cunningham Primary Intelligence Tests, Forms A and B, was 66.01 ± .39 in August, 1940, and after a year of pre-first grade training it was raised to 95.92 ± .27. The average IQ of the control group as obtained from the same tests was 68.90 ± .45 in August, 1940, and 78 ± .64 in the following year. The IQ's were almost parallel in 1940 and in 1941 there was a difference of 17.46 ± .64.

2) The children with a year of pre-first-grade work had an average percentile rank of 48.48 ± .40 on the norms of the Metropolitan Reading Readiness Tests as compared to an average of 7.81 ± .38 for the control group. The differences between the two averages was 40.67 ± .55. These percentile ranks were used as a basis of prediction for the attainment of a given grade status at the end of the first year, as presented in the Metropolitan Reading Readiness Test Manual. The chances of grade status for the end of the first year were as follows:

For the experimental group:

Sixteen per cent will reach the grade level of 1–9 or over
Fifty per cent will reach the grade level of 1–6 or over
Forty-seven per cent will fail to receive promotion

For the control group:

None of the children can expect promotion

3) The children of the experimental group made an average score of 46.32 ± .43 on the Metropolitan Reading Achievement Tests and the control group made an average score of 16.24 ± .75, with a difference of 20.08 ± .86 between the two groups. The grade placements attained from these test scores were as follows:

For the experimental group:

Twenty per cent had achieved a grade placement of 2–1 or over

Sixty-two per cent had achieved a grade placement of 1–9 or over

Four per cent were lower than 1–6 in grade placement

This shows that forty-six per cent more children than was predicted made a grade placement of 1–9 or higher and that all children were promoted to the second grade although forty-seven per cent had been predicted to fail promotion.

For the control group:

Four per cent attained the grade placement of 1–5

Eighty-eight per cent were at or below the grade placement of 1–3

4) The Metropolitan Reading Readiness Test and the Metropolitan Achievement Tests Primary I Battery; Form B, scores show a high positive correlation of .904 ± .008 between reading readiness and reading achievement. Therefore, those children who have had a preparatory year of work in order to develop reading readiness are at a great advantage because of their acquisition of social and emotional adjustments, and hence their possibility of success in reading.

5) The results from the Pintner-Cunningham Primary Intelligence Tests: Forms A and B, the Metropolitan Reading Readiness Tests, and the Metropolitan Achievement Tests, Primary I Battery: Form B signify that Spanish-American children with pre-first-grade training have a decided advantage over the children who do not have the experience of pre-first-grade training, and raises the presumption that a large percentage of failures

could be eliminated in the lower grades if all Spanish-American children were given a year of preparatory training.

6) It can be concluded from this study that pre-first-grade training is an important factor in success in learning to read among Spanish-American children, and that the above presented data justify the presentation of a pre-first-grade reading readiness program to children below the normal age of school entrance.

A TECHNIQUE FOR STUDYING THE BEHAVIOR OF MUSEUM VISITORS

L. C. NIELSEN*

The Laboratory School, University of Chicago

1. THE NEED FOR COMPREHENSIVE VISITOR STUDY

It is an interesting feature of the museum educational situation that although the activities of visitors are physically rather narrowly prescribed by aisles and exhibit fronts, the visitors themselves nevertheless have remained the least known, the most unpredictable, and the most difficult to study of all the aspects of museum work.

Technical problems of collecting, preserving, mounting, lighting, display, architecture, and even lecturing have been largely met and solved, *per se*. But these solutions may or may not be successful and meaningful when viewed from the standpoint of visitor reactions. Intended exhibit sequences may not be followed. Visitors may not stop long enough at each exhibit to derive any measure of meaning from it. The most significant exhibits may not receive the most attention.

The only method that can be employed to determine how museum visitors react is to observe the reactions of the visitor himself, in the context of specific museum situations, which may or may not be common to many museums. In other words, not only is it important to devise techniques leading to generalizations applicable to typical situations (such as identifying a general tendency to turn to the right), but also methods must be found whereby each museum can study its particular visitors in the light of its unique functions, gaining an intimate knowledge of their reactions to its distinctive techniques.

2. LIMITATIONS OF PAST APPROACHES

To the extent that museums have become aware of this problem, museum workers have not been attracted to the prospect of taking notebook and stopwatch in hand and trailing hundreds of visitors through the halls of their scientific or artistic storehouse

* Formerly Research Associate in Education, Museum of Science and Industry, Chicago, Illinois.

of culture. One need only turn to the records of the studies carried on by Robinson and Melton and their staffs to obtain an estimate of the labor which apparently lies in store for the progressive museum worker who wishes to undertake a study of the museum's constituency.

It is clear that although the trailing technique served admirably in the pioneer studies, it is not the final answer for any general or intensive program of visitor study. Aside from, or perhaps akin to, humanitarian considerations with respect to the museum worker, a second drawback to the traditional technique is its cost. If there is to be any widespread study by individual institutions, the research must fall within the scope of their financial capacities. The cost of training and maintaining one or more staff members for the exclusive purpose of trailing ten or twenty visitors a day until the hundreds of necessary cases have been accumulated would be a large item in most museum budgets.

Thirdly, there is the consideration of the reliability of these data once they are accumulated. Even an experienced observer will often feel that he has more than he can easily or comfortably handle as he attempts to time and characterize each stop-and-go made by his 'case,' adjust his stop-watch, jot down his data and keep the 'subject' in sight, while at the same time maintaining a discreet distance and an air of studied casualness.

These considerations in turn lead us to a fourth disadvantage of the trailing technique—lack of efficiency. By the time the observer has eliminated data for the 'cases' by whom he was detected or who for some reason he found impossible to record at all accurately, a day's work is not likely to yield him more than twenty-odd complete records. In addition, the number of aspects of visitor behavior that an observer can record in one 'round' is relatively limited. It would be virtually impossible for him consistently to transcribe path, time of stay, labels read, degree of physical participation in the exhibit (if operative), and closeness or survey-character of the visitors' observation.

3. PROPOSAL OF A NEW TECHNIQUE

In the hope that it might prove useful to those interested in determining what visitors do in any given room or exhibit of their museum, this article is intended as a report of an alternative

technique of visitor observation which is proving quite successful in the Museum of Science and Industry in Chicago.

The procedure is essentially one of time-lapse photography. A motion-picture camera equipped with single-exposure release is mounted high in one corner of the room, concealed as much as possible from visitor view. By means of a simple motor drive and appropriate arrangement of cams, the single-exposure release may be tripped at any pre-determined interval of time, and pictures taken correspondingly at any desired rate, though a range of from one to sixty frames per minute is the most generally useful. The resulting film-strip, a series of snapshots of the entire room, gives a truly candid, 'slow-motion' record of all that occurred in the room during the photographic period—how many visitors entered the room, or how many there were at any given moment; where they went and in what sequence; how they looked at the exhibits; how long they stayed at each exhibit or in the room as a whole.

a. Implementation of the Photographic Technique

This procedure has been used in the Museum of Science and Industry in connection with visitor study in an exhibit designed to tell the Story of Magnetism (see Fig. 1). The room in question is small, only thirty feet square; yet the smallest angle which will include all the public area (aisles and exhibit-fronts) is about 70°. The average movie camera has a field only slightly better than 20°. At the time the technique was first implemented, war conditions had rendered wide-angle lenses with which to experiment completely unavailable. Even now a lense capable of covering this wide an angle without appreciably cutting down speed is difficult to obtain. The answer to the problem of compressing the 70° field occupied by the room into the 20° field covered by the camera was found through the use of a large convex mirror (aperture fifteen inches, radius of curvature thirty inches), the camera being pointed directly at the mirror, and focused on the virtual image of the room formed by and behind the mirror surface. The slight distortion of image caused by this arrangement offered no real handicap in the interpretation or analysis of the film.

The trial runs were made on 35 mm ultra-speed film (the camera and timing mechanism for the purpose were kindly lent by

Mr. Herbert A. Thelen of the University of Chicago Department of Education), with a #2 photoflood in the regular central overhead installation, and the regular fluorescent fixtures illuminating the exhibits. The intervals between exposures were varied from five to fifteen seconds. For this particular exhibit, it was found that seven and one-half seconds was the most useful interval, not too slow to catch significant transitions of

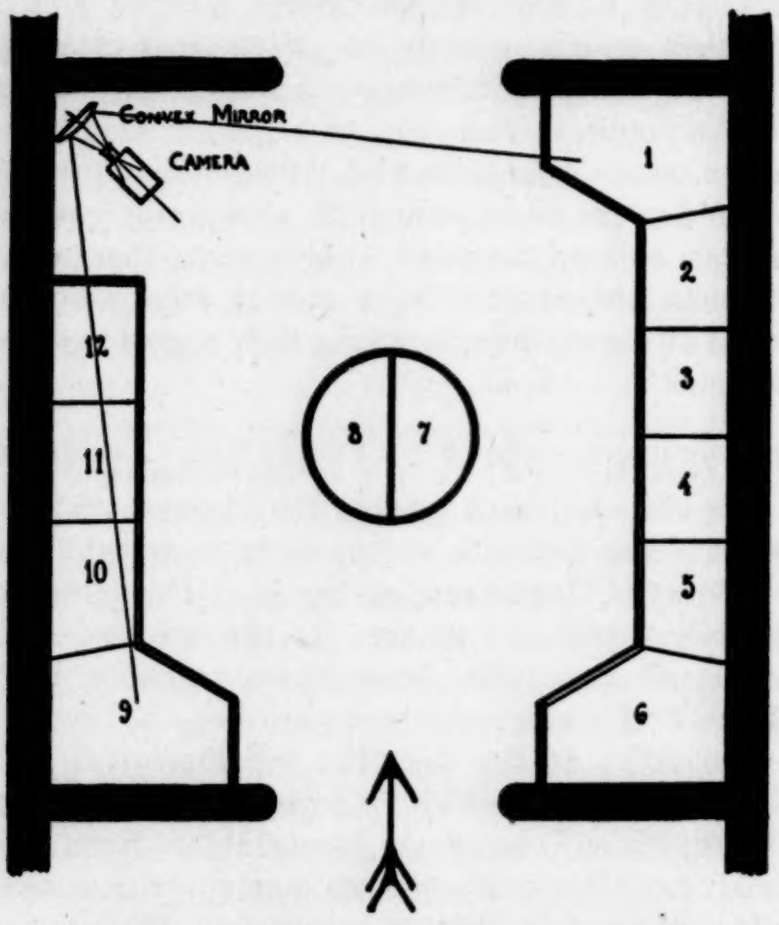

FIG. 1.—Diagrammatic representation of photographic arrangement in exhibit.

visitors from one exhibit to another, nor fast enough to cause successive pictures to be nearly identical.

The seven and one-half second interval causes 35 mm film to be used at the rate of thirty feet per hour. A 16 mm film, though not as desirable, is satisfactory, and is consumed at the rate of only twelve feet per hour. In either case, however, the cost cannot be compared to that of an observer working an equal length of time on a small fraction of the number of cases that the film is able to record completely.

b. Transcription of Film Record

Though by no means exhaustive of the possibilities, discussion of the type of analyses so far made with these film records might be of interest, as an indication of the versatility and wide range of usefulness of this technique.

Since it was anticipated that the data on the film would be 'gone over' from a number of different standpoints and for different purposes, the first step was to eliminate the necessity for continuous and repeated use of projection encumbrances. This was done by 'transcribing' the film record onto tabular form, as follows (see Fig. 2):

Each horizontal line of the square-ruled sheet is used to represent one picture, or frame, of the film strip. Each vertical column represents a specific location in the room. Any person identified at a particular location on a particular exposure is assigned a tally in the appropriate square of the sheet. If the person is a man, this tally takes somewhat the form of a '9'; if a woman, '6'; if a child, a small 'i.' Tots under four years of age are not counted. If a given person appears on one picture only, he receives the corresponding tally and that is all. But if this person is identified through several successive frames, then the tallies made for him are all connected by a line, so that his path vertically along the sheet corresponds to his path in the room.

c. Treatment of Data

Several types of analysis immediately suggest themselves. Mere inspection of this record reveals much concerning the general paths followed and population densities in various parts of the room.

Frequency distributions for the times of stay at each exhibit may easily be made, which in turn enable the determination of the mean, median, and modal observation time for each exhibit or for the room as a whole. These figures in turn may be interpreted as indications of the attractiveness, difficulty, 'drawing-power,' or interest of the various exhibits. In addition, an extremely powerful and suggestive device is that of identifying quartiles or centiles of observation time, through which the museum educator is provided with a basis of estimating what

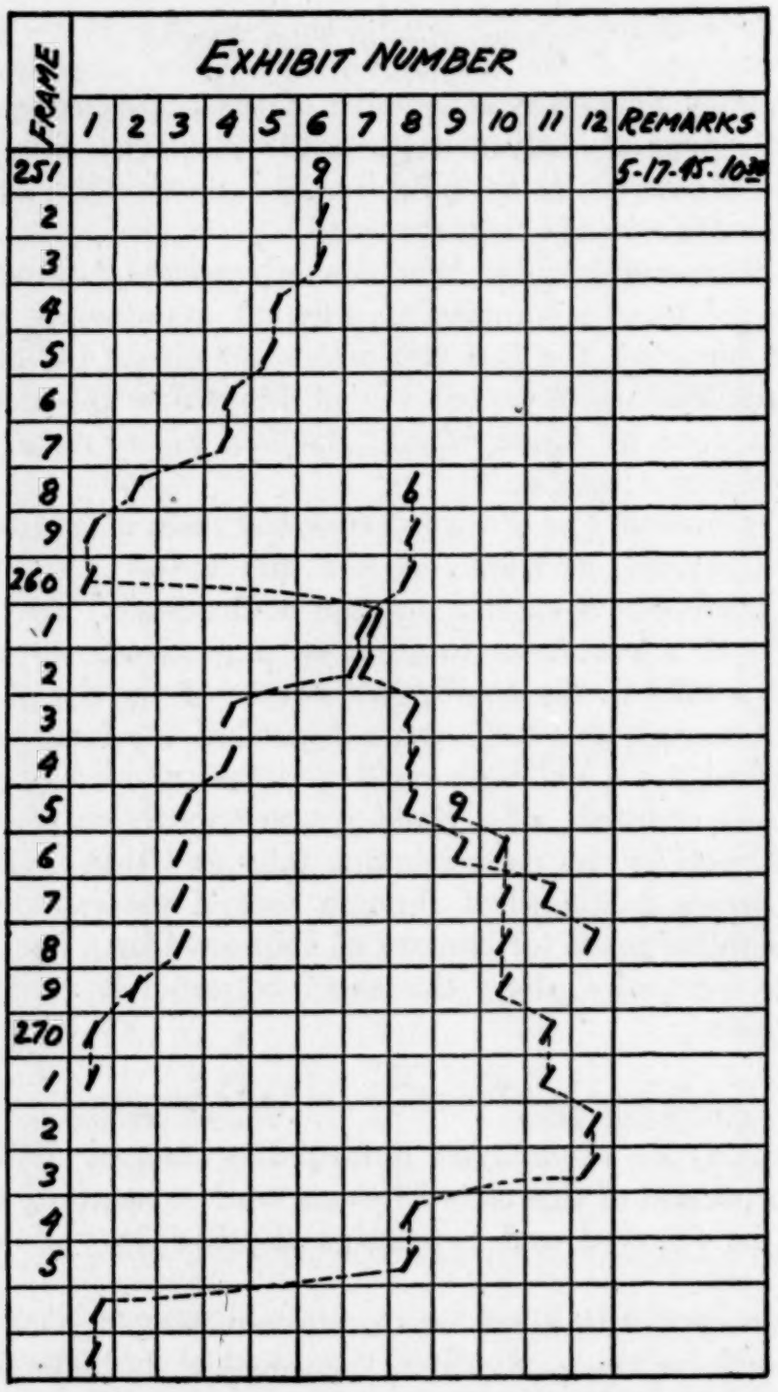

FIG. 2.—Sample transcription of visitor paths from photographic record.

he may expect from different proportions of the public in the way of attention and time spent.

The various paths identified may also be summarized and classified, and generalizations made concerning routing. Further experiments at guidance may be carried out and evaluated by

this technique. Special analyses might also be made of children, or of the behavior of men *versus* women, accompanied *versus* lone, summer compared to winter, or morning compared to afternoon visitors.

d. Validity and Limitations of the Technique

The analyses and types of inferences indicated above have actually been carried out for the Magnetism Exhibit of the Museum.

The routing studies corroborated subjective impressions that visitors were not seeing the exhibit in proper sequence. Further photographic studies with various experimental methods of guidance provided quantitative bases for comparing the effectiveness of these methods.

The statistical analysis of time-of-stay-at-each-exhibit, revealing as it did variations in average time of from nine seconds to one minute, and ranges of from five seconds to three and one-half minutes per exhibit holds great educational implications, which will be reported upon in a later paper.

The two chief factors influencing the validity of this method of visitor study are accuracy and care of photographic technique, and compromising at the right point between recording significant visitor motions and making efficient use of materials and analyzing time. A sixteen-frame-per-second motion-picture record would catch every movement made by a visitor; if head and hand motions, for example, were considered to be significant aspects of visitor behavior for a particular study, then this speed might be an appropriate one. However, if only major transitions between exhibits are desired for purposes of study, the frequency of exposure can be decreased according to the characteristics of the public, and the discretion of the experimenter.

Practical limitations of photography for museum purposes are felt mainly in connection with size, crowding, and illumination of the room. The larger the room, the more difficult it becomes to include all of it in the field of the camera. For this, only wide-angle lenses, convex mirrors, or coördinate use of two cameras can be suggested. When crowds become very dense, visitors tend to obscure one another in perspective, and it is difficult to identify and trace single visitors along the film record. A good lense and use of 35 mm film for good resolution, and a high

camera angle are good precautions in this respect, though it is suggested that very crowded periods be avoided for traffic studies, unless information concerning such situations is especially desired. The problems of illumination are parallel to those met in any indoor motion-picture photographic work; it will usually be found necessary to supplement the regular lighting installations with some type of floodlight.

CONCLUSIONS

It is necessary to recognize that picture-analysis alone does not yield a total picture of the visitor, but only of his behavior in specific situations. Supplementing the picture record with sound recording would provide still more complete information concerning visitor reactions. But such questions as what factors prompted the visitor to come, what was his emotional set or state of fatigue at the time of his visit, and so forth, require other techniques for their answer.

Such factors, however, though interesting, are extremely variable and erratic, and—what is more important—not subject to museum control. Study of them, except for ascertaining the educational backgrounds and interests of the visitors, are not likely to yield much that is useful as a basis for museum action.

The museum must take its visitors as they are—tired, hurried, or uninterested—and seek ways of making their visits, short or long, as fruitful educationally as possible. Studying visitor reactions to the educational experiences within the museum is a necessary first step in evaluating the extent of fruition of museum educational efforts, and a necessary link between present practice and even better, more effective, practice.

AN EXPERIMENT IN SELF-ANALYSIS

GEORGE S. BEERY

State Teachers College, Whitewater, Wis.

In the Orientation classes of Whitewater State Teachers College, one of the standard practices is to give two tests of mental ability during the year. This practice is prescribed by the catalogue. However, until this year not a great deal of follow-up work was done with respect to the results of these intelligence tests. Other tests had also been given at intervals, particularly reading tests, personality tests, study habits inventory, and vocational inventories. Unfortunately, many of the results of these tests were never used for serious and organized guidance purposes.

For some years the writer has been collecting comments from students regarding their experiences with tests. These reactions were catalogued for the most part from juniors and formed a background for certain work in Educational Tests and Measurements. The reactions which appeared on many papers were somewhat as follows: "Since I never knew the results of my tests, I never thought that they were particularly useful." "Our teacher refused to let us see the results of our tests, saying that it would be bad for us to know the results." "Test results were always treated with a certain degree of mystery." "I took standardized tests as a matter of course and never thought any more about them." "To my knowledge, the purpose and results of the tests were never disclosed to the student."

These reactions were somewhat at odds with comments made in books on psychology and educational measurements to the effect that perhaps students should not be informed of the results of their tests. There was a theory that more harm than good would come out of revealing to students the results of their tests. Such an attitude with respect to the giving out of test results may have resulted in much of the student dislike for standardized tests.

The classes in Orientation were approached this year with a feeling that a lot of good might come out of coöperative guidance experiences. The course in Orientation was designed according to the catalogue to give "guidance in adjustments to college life with emphasis upon familiarizing the student with his new

environment, its challenges and requirements, its mechanical and personal elements." Certainly the guidance could not be done wholly from the outside, but had to include the complete coöperation of the student in making himself amenable to guidance practices. During the first part of the year, the following standardized tests were administered to the college freshmen: American Council Psychological Examination, The Iowa Silent Reading Test, and the Bell Adjustment Inventory. The Otis Quick Scoring Test of Mental Ability was given at the beginning of the second semester. Although each of the students could have been called in and told his test results in a personal conference, it was decided to experiment with group guidance and to familiarize the student with the results of his tests and to begin with the group explanation before calling students in for individual conferences relative to their potentialities as students.

A personal analysis sheet was designed according to the following principles:

1) The form must be readily understandable.

2) The test results must be coördinated and must be interpreted graphically.

3) The student must be able to find out 'what he can do.'

4) A student should be able to see how he is adjusted as a person.

5) The student should be able to compare 'what he can do' with 'what he is doing.'

6) The scheme of presentation of test results and of other guidance material must be consistent.

Accordingly, a Personal Analysis Sheet was developed in an attempt to meet the various principles set up for a guidance device of this sort.

The results of the tests were given in terms of percentile ranks because these seemed to be the most easily understandable. The results of the various measures of 'what I can do' were connected so as to form a profile. This profile gave a graphic picture of the capabilities as determined by the tests listed. The results on the Bell Adjustment Inventory were also connected to form a 'personality profile.'

The student recorded at the bottom of the page his various subjects and indicated his grades at the end of nine weeks and—

Whitewater State Teachers College
PERSONAL ANALYSIS SHEET
Freshmen—1944–45

Name_____ Age____ Date_____
(Last name first)

(Percentile Ranks)

	0	10	20	30	40	50	60	70	80	90	100
Per. Rank in Graduating Class											
Henmon-Nelson Per. Rank											
Ave. Per. Rank (See Cat. p. 14)											
Am. Coun. Psych. Exam.											
Iowa Silent Reading: Comprehension											
Rate											
Otis Quick-Scoring Mental Ability Test											

(Bell Adjustment Inventory Ratings)

	Very unsat.	Unsat.	Average	Good	Excellent
Bell Adjustment Inventory:					
Home					
Health					
Social					
Emotional					

Grades

Name of Subject	E	D	C	B	A
1					
2					
3					
4					
5					
6					
7					
8					
9					

at the end of the semester—his semester grades. Similarly, these grades were connected so as to form a profile of 'what I am doing.' The Personal Analysis Sheet was prepared in duplicate so that the student could have a copy; the other copy going to the office of the Director of Personnel.

Instructions were prepared to accompany the Personal Analysis Sheet so that the student could read about and interpret the various items. The instructions were as follows:

(———)

"The purpose of the Personal Analysis Sheet is to lay before you certain measures of your ability, both general mental ability and reading ability, in addition to the personality measures as given on the Bell Adjustment Inventory. These measures of ability are then to be compared with the measures of actual accomplishment as indicated by the grades which you have earned.

"The percentile rank in the graduating class is an indication of the per cent of your graduating class that was below you in accomplishment. For example, if you graduated eleventh in a class of fifty-two your percentile rank was 78. Your rank of 11 was better than that of the remaining forty-one members of your class. (Forty-one is seventy-eight per cent of fifty-two.) In this way we are able to compare the percentile ranks of members of graduating classes of unequal size.

"Your percentile rank on the Henmon-Nelson Test indicates the position of your score with relation to some thirty or forty thousand high-school seniors in Wisconsin high schools who took the Henmon-Nelson Test of Mental Ability in the senior year in high school. For example, a percentile rank of 73 on the Henmon-Nelson test indicates that your score on that test was better than the score earned by seventy-three of the total number of seniors who took the test.

"The average percentile rank is a combination of the percentile rank on the Henmon-Nelson Test of Mental Ability and the percentile rank in the graduating class, according to the Regent's formula.[1] For example, the formula calls for two times the

[1] The catalogue statement explaining this particular procedure is as follows:

C. Scholastic Requirements for Admission. By action of the Board of Regents of Normal Schools, Resolution 695, scholastic requirement stand-

percentile rank in the graduating class plus the percentile rank on the Henmon-Nelson Test, the sum to be divided by 3. An example of this would be as follows: The percentile rank in the graduating class being 78 and the percentile rank on the Henmon-Nelson Test being 73, take 2 × 78 or 156; add 73 which gives 229; divide by 3 and the result is 76 which is the average percentile rank. This is the score which is taken as a basis for admission to college. The college catalogue states that a minimum average percentile rank of 30 is necessary for success in college work as well as for admission to college.

"In a similar fashion the results of the mental test which you took earlier in the year—The American Council Psychological Examination—are interpreted in terms of percentile rank. A percentile rank here of 70 means that your score was better than seventy per cent of the scores of the local freshman class. The scores for all of the freshmen over the country who took the American Council Psychological Examination are not yet available, so it has been necessary to interpret the results of this examination in terms of the local freshman class. When the national percentile ranks become available they will be, no doubt, slightly higher than the percentile ranks as given on the sheet.

"In interpreting the results on the mental ability tests it is well to remember that in most cases it is impossible for you to do better than you can on a mental ability test; whereas it is pos-

---

ards based upon rank in high-school class and percentile rank on the Wisconsin Coöperative Aptitude Test were set up for admission to the State Teachers Colleges in Wisconsin beginning September, 1940. The following formula is used to determine the average percentile rank upon which the admission of the student is based:

$$APR = \frac{2\,HS + \text{Test}}{3}$$

The interpretation of this formula is as follows: "HS" stands for the percentile rank in high-school graduating class. "Test" indicates the percentile rank on the Wisconsin Coöperative Aptitude Test. "APR" equals the average percentile rank by formula. By action of the Board of Regents, students who have an APR of less than 30 do not meet the scholastic requirements for admission. [*Bulletin of the Whitewater State Teachers College*, Whitewater, Wisconsin, July 1944, General Catalogue, Serial No. 265.]

sible to do less well than you are actually able to do. The way you feel, the particular emotional state in which you are, or certain circumstances or environment may bring down your score so that on one test you may do less well than you did on another. However, if you strike an average between your percentile rank on the Henmon-Nelson Test and on the American Council Test you have a reasonably satisfactory estimate of your mental ability as indicated on the two tests. We understand that tests do not tell all the story as far as our mental ability is concerned, but they do give us about as good a measurement as we are able to get at the present time.

"On the Iowa Silent Reading Test we measure your ability in reading. The percentile rank in comprehension indicates your ability to understand what you read; whereas the percentile rank on rate is a measure of your comparative speed of reading. The percentile ranks on the Iowa Test are in terms of thousands of college freshmen.

"On the Bell Adjustment Inventory five areas of adjustment are indicated: Excellent, Good, Average, Unsatisfactory, and Very Unsatisfactory. These are in comparison with all college men and women who took the test. Again, it must be understood that the Adjustment Inventory is an attempt to find out how you feel or how you make adjustments in the areas indicated: Home, Health, Social Adjustment, and Emotional Adjustment.

"For the most part your grades as earned at the end of the nine weeks may be interpreted thus: A, excellent; B, good; C, average; D, unsatisfactory; E, very unsatisfactory. By comparing the grades which you earned in your various subjects with the figures at the top of the page which indicate your general mental ability and your reading ability, you will have some idea of whether you are meeting what is expected of you or whether you are falling below what might be expected of you. If you do less well than is expected of you, perhaps the results on the Bell Adjustment Inventory will give you a clue to your lack of adjustment; or perhaps you may re-examine your study habits to see if there is a reason why you are not earning the grades that you should have. There may also be other personal factors involved which only you, yourself, know about.

"If you have any questions on the Personal Analysis Sheet will you please write them down, sign your name to them, and

leave them in my office so that as soon as possible we can have a conference on your problem?"

(———)

Comparable descriptions of the items included on the Personal Analysis Sheet were as follows:

Percentile Ranks	Bell Adjustment Inventory Ratings	Grades
80–100	Excellent	A
60–80	Good	B
40–60	Average	C
20–40	Unsatisfactory	D
0–20	Very Unsatisfactory	E

Thus, it was rather easy for the student to glance at the profile in the 'what I can do area' and compare it with himself as a person as found in the ratings on the Bell Adjustment Inventory and with the 'what I am doing' profile as found at the bottom of the page on his subject grades.

As a subjective check upon the student reaction to the Personal Analysis Sheet, a reaction sheet was issued to the students as follows:

Whitewater State Teachers College

REACTIONS TO PERSONAL ANALYSIS SHEET

Directions: Check the statement which most nearly expresses your reaction to the particular item indicated.

I. What is your reaction to the items indicating your ability?

A. Percentile rank in graduating class:

* 5% 1. I don't understand it.

2% 2. The percentile rank is wrong.

62% 3. I agree with it.

25% 4. No comment.

6% 5. Remarks:_____

* These figures represent the percent of the group giving the reactions marked. A total of 107 students filed complete Personal Reaction Sheets and the percents given are based upon this number. The raw numbers marking each blank have been omitted in favor of the number representing the per cent of the group.

B. Percentile rank on Henmon-Nelson test:

4% 1. I don't understand it.
15% 2. It seems too low.
2% 3. It seems too high.
0 4. I resent this.
61% 5. No comment.
18% 6. Remarks:_____

C. The American Council Psychological Examination:

0 1. The test is unfair.
26% 2. The results seem about right.
28% 3. The results seem too low.
3% 4. I resent this.
32% 5. No comment.
11% 6. Remarks:_____

D. The Iowa Silent Reading test:

27% 1. The test does not indicate my true reading ability.
4% 2. The test was unfair.
57% 3. The results seem about right.
1% 4. I don't need to know how well I can read.
11% 5. Remarks:_____

II. What is your reaction to the items on the Personality Inventory?

	Adjustments: a. Home	b. Health	c. Social	d. Emotional
1. I can't agree with this.	24%	18%	12%	15%
2. The test results are correct.	53%	52%	54%	51%
3. I resent the test results.	0	0	2%	5%
4. No comment.	13%	22%	21%	18%
5. Remarks:_____	10%	8%	11%	11%

III. What is your reaction to your grades:

30% 1. They are too low.
0 2. I did better than my grades indicate.
0 3. I was marked unfairly.
38% 4. The grades are o.k.
12% 5. I worked up to my capacity.
20% 6. Remarks:_____

IV. What is your reaction to the whole sheet:

1% 1. It is unfair.
77% 2. It revealed things I never knew before.
11% 3. I don't like it.
1% 4. Such things should not be shown to the student.
10% 5. Remarks:_____

There was a surprisingly small number of written remarks on the reaction sheets. Most of the students merely checked the appropriate statements provided. Typical written reactions in the various areas are as follows:

I. A. What is your reaction to the percentile rank in the graduating class?

1) I know I didn't work nearly as much as I should have in high school.
2) I should have been higher.
3) I knew my rank before I came to college.
4) It seems some students who did poorer work from day to day came out ahead of me in my class rank.
5) Should have worked harder in high school.
6) Not true measure of achievement of ability of myself or classmates—not enough factors considered.
7) I was on the Honor Roll and the rank seems rather low for that.
8) I never knew what my rank in graduating class was.

B. What is your reaction to the percentile rank on the Henmon-Nelson test?

1) Seems o.k.

2) O.K. Kinda forget about it.
3) I think it is correct.
4) I didn't think it counted much.
5) Did not think it was so important.
6) Importance of it was not stressed.
7) Why is my reading ability good and my Henmon-Nelson test low?
8) I don't know what this test means.
9) It seems too low, but I think it's right. It seems a bit strange to have my IQ so much lower than my rank in my graduating class. My reading ability is low, but I expected that to be the case. This inability didn't hinder me much in high school but it does quite a bit now in college. What to do?

C. What is your reaction to the American Council Psychological Examination?

1) I do not believe I took the test under proper conditions.
2) According to my other tests it doesn't seem to balance.
3) Averaging the first two tests, this mark doesn't seem logical.
4) Would like to find out score in specific cases: arithmetic, spelling; in fact, every part.
5) Can't figure out what happened there.
6) I believe I could have done better.
7) It bothers me to be so low on this.
8) It seems to me my results should be higher.
9) I think I did my best that day.

D. What is your reaction to the Iowa Silent Reading Test?

1) I know that I am a slow reader.
2) I believe I can read better than the tests indicate.
3) I didn't think my reading ability was this low.
4) The test revealed my very slow rate of reading which I wasn't entirely aware of.
5) I can read better when not tested.
6) I know I am a slow reader, but I do not like to read.
7) In reading my assignments it seems I read faster and comprehend better.
8) Can read as fast as others, but tests prove differently.

9) In all other reading tests I have always read nearly one hundred per cent so for this reason something is wrong with the test.

E. What is your reaction to the items on the Personality Inventory?

1) I believe my home to be as good as the average. I didn't know I was emotionally unstable.
2) I can't understand being emotionally unsatisfactory.
3) All these results were below average which is very odd.
4) I think it's o.k.
5) I don't think it is a good way of getting information.
6) When writing that I checked just anything not knowing they would count.
7) As a whole I feel that I have adjusted better than it indicates.
8) My social is very unsatisfactory—what to do?
9) I think I know myself better.

F. What is your reaction to your grades?

1) I worked fairly hard but know I could do better.
2) I'm not too worried about my grades since I try my best.
3) In one subject I think I did better than what my grades showed.
4) They are too low, but it was the best I could do!
5) I did not work up to my capacity. I think I could do better.
6) My own fault; should work harder.
7) I seem to be working all the time, but my grades don't show it. Why?
8) I would like to attain better grades, but how?
9) The grades are o.k., but I expected to be able to do better.
10) I haven't worked as hard as I should so I can't expect any better grades.
11) Guess I'd better 'pitch in' in the future.
12) Not the teacher's fault, but my own.
13) I think I'm getting myself adjusted and during my second semester I hope to do better.
14) Would have liked something higher, but did not work for it. Feel it's possible to do better; lacked desire.

G. What is your reaction to the whole sheet?

1) I think it is a good thing for a student to see, because then you know.
2) It may be fun to get ahead of the mental capacity.
3) Shows exactly what can be done, and it is somewhat of a challenge.
4) Tells what I wanted to know.
5) All high-school seniors, college students should know where they stand.
6) I still don't know what's wrong.
7) I think every student should be shown his or her results on analysis tests.
8) It is interesting to know whether bad or otherwise.
9) Very glad to get it.
10) I think it is a good thing—shows a student where he stands.
11) Everyone may not have been up to par the time the tests were given!
12) It shows me how much better I could do if I wanted to.
13) It shows me just how dumb I really am!
14) It caused me a lot of trouble; and I don't agree with it in all places.
15) May prove a stimulant to do better.
16) A good idea—now I know where I stand!
17) I think it is a good idea to see what we need to do.
18) I don't like it because it reveals to me definitely facts that I was afraid of being, so I know I shouldn't mind knowing these things, but I do.

It is interesting to note that the largest per cent of the students agreed with their standing. Most of the students seem to accept their rating on the Henmon-Nelson Test (this test is given to all seniors in Wisconsin high schools and the results are given to the teachers colleges and high schools).

The reactions to the American Council Psychological Examination ('results seem too low') appears to indicate in this particular case either: (1) The test was not properly administered, a fact which would result in inadequate and incorrect results; or (2) The American Council Psychological Examination was not sufficiently well explained to the students so that they could

sense the true meaning of the results or the high regard which educators have for the test.

The reactions on the Iowa Silent Reading Test seem to indicate that quite a few students were not familiar with reading tests and their purpose. They, perhaps, did not know that it is possible to secure a measure of a person's reading rate and comprehension by means of a formal reading test. Their unfamiliarity with reading tests may also have given rise to this particular reaction.

The greatest reaction of a negative sort was expected on the personality inventory. However, strangely enough, such negative reaction, by and large, did not materialize in as great a quantity as was expected. The number of students who 'resented the test results' was also surprisingly small. Many of the students who didn't agree with the test indications put the cause of their disagreement in remarks, and these remarks have already been cited. Perhaps the most pertinent remark was one to the effect: "I know that I have an unsatisfactory social adjustment; what can I do about it?" This reaction in itself indicated the need for further guidance.

A reasonable number of students considered their grades too low, but no one was disposed to say that "he did better than his grades indicates" or that "he was marked unfairly." The number who 'worked up to capacity' was rather interesting. There was a larger number of written comments and reactions to grades than in any other area. Many of the reactions were typified by the following: "Guess I had better pitch in in the future." "Some subjects I can't seem to do good on." (Apparently this student needs some instruction in English also!)

The general reactions to the whole sheet were quite interesting. For the most part, students realized that they were finding out something they never knew before. There was a surprisingly small number of them who definitely disliked it. Only one student said it was unfair, and another said that "such things should not be known to the student."

The "longtime" results of this experiment were tried out during the second semester. Frequent references were made to the Personal Analysis Sheets. Conferences were held with the individual students in an effort to detect the particular areas in which considerable guidance needed to be done for the benefit

of a student. Explanations of the marks they received were based on the profiles from the Personal Analysis Sheets. Material studied in the area of occupational information and vocational guidance—at which time personal self-analysis was stressed as a necessary step in the selection of a vocation—was made much more pertinent and of direct benefit by reason of the student having at his fingertips visual records of what he could do and what he was doing and what sort of a person he was.

Though this experiment in self-analysis is not to be taken as the 'last word' in such techniques, it may serve as a guide for future work in the very necessary area of self-analysis.

BOOK REVIEWS

ARNOLD GESELL, FRANCES L. ILG AND OTHERS. *Infant and Child in the Culture of Today.* New York: The Macmillan Co., pp. 399. $4.00

This remarkable book describes in clear, informal language the child's development at twelve levels from birth through age five. The scope of the book is indicated by sixty-five pictures assembled in the end-papers to show the progression of behavior during the first five years of life.

'Behavior profiles' are employed to summarize the most typical traits and growth trends for each level. These profiles are filled with authoritative facts, penetrating insights, and practical suggestions for child guidance. Following each 'behavior profile' is a section entitled "Behavior Day" which sets forth details of child care under headings such as sleep, feeding, elimination, and so forth. Although the significance of individual variation is recognized and stressed, the reader is warned that age norms are not set up as standards, and that care must be used in all comparisons. The norms are designed only for purposes of orientation and interpretation.

The twelve behavior profiles, read in order, reveal the flow and essential continuity of human development; yet each profile stands alone—a self-contained, independent description of a single level of human growth. These descriptions afford useful bases on which to build sound guidance practices; but they have an even larger significance since "the growth characteristics of the early years—must determine our whole outlook upon the nature and needs of the individual child. These growth characteristics are so fundamental that not even the most modern culture can supersede them." In fact, a culture developed in accordance with these facts might lead to the highest expression of democracy.

The household is viewed as a cultural workshop in which the democratic way of life is fostered through a respect for the unique nature and needs of each child. It is the place where human relationships have their origin and respect for individuality begins. Accordingly, unusual significance is attached to the first year of life. After the child is eighteen months of age, "his behavior extends beyond the confines of the home. He goes abroad." This excursion may take him to the nursery school.

In any event, it will lead to broader contacts and social problems for which psychological care is needed. The contribution of one agency—the nursery school—is fully discussed. However, the treatment covering the preschool ages is presented in such a way as to be serviceable in the guidance of young children whether or not they attend nursery school.

A 'developmental philosophy' is found to pervade the book. At no time in history was this philosophy so urgently needed as it is today, since "no previous culture has ever achieved a product more magnificent than the present body of natural and engineering science. This achievement is our hope as well as our despair. The despair will not lessen until the techniques of modern science can be more sincerely brought to bear on the problem of behavior." Since it is impossible to understand conduct and behavior at any age level without a consideration of antecedents, it is imperative that all adults concerned with the care of boys and girls obtain adequate knowledge of the growth of infants and young children. For it is in the early years that many significant patterns appear and are shaped. Because of these facts, "the professional training of school teachers should include a liberalizing acquaintance with the developmental psychology of infancy to offset the stilted textbook limitations of an educational psychology too narrowly based on a study of the learning process." Moreover, it is particularly fitting that American teachers acquire a developmental philosophy since "America, more than any other country in the world, has dedicated its scientific resources and good will to the furtherance of the free development of her children."

Attention should be directed to an outstanding feature of this book. The Appendix contains lists of play materials for the different levels, and carefully selected books, recordings, and other indispensable materials for child care and guidance.

PAUL WITTY

Northwestern University

ROY M. DORCUS AND G. WILSON SHAFFER. *Textbook of Abnormal Psychology.* Baltimore: The Williams & Wilkins Company, 1945, pp. 547.

Like its predecessors, the third edition of this frequently used textbook in abnormal psychology is aimed at satisfying the needs

of advanced students in psychology and premedical or medical students. The authors consciously ignore the problems of adapting the text to those students whose preparation in psychology or related sciences is inadequate. However, this volume like its preceding editions is relatively well written, and instructors who have been satisfied with the previous editions will be satisfied with this one.

The authors talk of many changes they have in their literary references and in researches reported. However, a close investigation shows that, by and large, the framework as well as the general attitude and treatment in the second and third editions are very much alike; and the content, for the larger portion, is also alike.

There are, however, some changes which deserve mention. For example, the consideration of abnormal psychology and sociology is somewhat expanded with more attention given to the researches of Landis and Page, particularly the point emphasized by these researches to the effect that incidence of mental disorder is related to bed capacity and facilities of hospitals rather than to war and depression.

The authors' attitude towards the validity of fixed symbolism in dream interpretation remains pretty much the same. However, whereas in the second volume the denial or the invalidity is expressed with the help of a gossipy story, in the third edition this is replaced by a quotation from Sears' evaluation of psychoanalytic concepts. Their treatment of experimental studies of memory and feeling is somewhat more adequate but still omits the central connection in these researches as, for example, portrayed by Rapaport in his recent volume on memory and emotion. In their chapter on "Treatment by Physical and Chemical Assault" is included a consideration of psychosurgery or prefrontal lobotomy, narcosis therapy, benzedrine sulphate therapy, and vitamin therapy. A word and a phrase here and there indicate that the authors are aware of increasing shifts towards a dynamic interpretation in psychiatry as well as in clinical psychology. However, not sufficient attention is given to some of the more significant researches in the psychodynamics of behavior. For example, Rorschach researches and projected techniques of personality evaluation are given little consideration, and some of the concepts as well as researches reported at length in Maslow

and Mittlemann's abnormal psychology textbook are considered very little in this volume. In fact, for graduate students it is the impression of the reviewer that supplementation by a book like that of Maslow and Mittleman's would be advisable.

All in all, then, we have in this text a very sound covering of the literature with a relative lack of bias. Its weakest stages are in researches on dynamics of human nature as well as in therapy. Omitted also are the researches on frustration by Kurt Lewin and his students. However, it could serve as a very satisfactory textbook for a comprehensive course in abnormal psychology.

H. MELTZER

Psychological Service Center, St. Louis, Missouri